




[Center for Machine Learning and Intelligent Systems](#)

[About](#) [Citation Policy](#) [Donate a Data Set](#) [Contact](#)

☒ Repository
 ☐ Web
 

[View ALL Data Sets](#)

Flags Data Set

Download: [Data Folder](#), [Data Set Description](#)

Abstract: From Collins Gem Guide to Flags, 1986



Data Set Characteristics:	Multivariate	Number of Instances:	194	Area:	N/A
Attribute Characteristics:	Categorical, Integer	Number of Attributes:	30	Date Donated	1990-05-15
Associated Tasks:	Classification	Missing Values?	No	Number of Web Hits:	101853

Source:

Creators:

Collected primarily from the "Collins Gem Guide to Flags": Collins Publishers (1986).

Donor:

Richard S. Forsyth
8 Grosvenor Avenue
Mapperley Park
Nottingham NG3 5DX
0602-621676

Data Set Information:

This data file contains details of various nations and their flags. In this file the fields are separated by spaces (not commas). With this data you can try things like predicting the religion of a country from its size and the colours in its flag.

10 attributes are numeric-valued. The remainder are either Boolean- or nominal-valued.

Attribute Information:

1. name: Name of the country concerned
2. landmass: 1=N.America, 2=S.America, 3=Europe, 4=Africa, 4=Asia, 6=Oceania
3. zone: Geographic quadrant, based on Greenwich and the Equator; 1=NE, 2=SE, 3=SW, 4=NW
4. area: in thousands of square km
5. population: in round millions
6. language: 1=English, 2=Spanish, 3=French, 4=German, 5=Slavic, 6=Other Indo-European, 7=Chinese, 8=Arabic, 9=Japanese/Turkish/Finnish/Magyar, 10=Others
7. religion: 0=Catholic, 1=Other Christian, 2=Muslim, 3=Buddhist, 4=Hindu, 5=Ethnic, 6=Marxist, 7=Others
8. bars: Number of vertical bars in the flag
9. stripes: Number of horizontal stripes in the flag
10. colours: Number of different colours in the flag
11. red: 0 if red absent, 1 if red present in the flag
12. green: same for green
13. blue: same for blue
14. gold: same for gold (also yellow)
15. white: same for white
16. black: same for black
17. orange: same for orange (also brown)
18. mainhue: predominant colour in the flag (tie-breaks decided by taking the topmost hue, if that fails then the most central hue, and if that fails the leftmost hue)
19. circles: Number of circles in the flag
20. crosses: Number of (upright) crosses
21. saltires: Number of diagonal crosses
22. quarters: Number of quartered sections
23. sunstars: Number of sun or star symbols
24. crescent: 1 if a crescent moon symbol present, else 0
25. triangle: 1 if any triangles present, 0 otherwise
26. icon: 1 if an inanimate image present (e.g., a boat), otherwise 0
27. animate: 1 if an animate image (e.g., an eagle, a tree, a human hand) present, 0 otherwise
28. text: 1 if any letters or writing on the flag (e.g., a motto or slogan), 0 otherwise
29. topleft: colour in the top-left corner (moving right to decide tie-breaks)
30. botright: Colour in the bottom-left corner (moving left to decide tie-breaks)

Relevant Papers:

Forsyth's PC/BEAGLE User's Guide.

Papers That Cite This Data Set¹:



George H. John and Ron Kohavi and Karl Pfleger. Irrelevant Features and the Subset Selection Problem. ICML. 1994. [[View Context](#)].

Ron Kohavi and George H. John and Richard Long and David Manley and Karl Pfleger. MLC++: A Machine Learning Library in C. ICTAI. 1994. [[View Context](#)].

Wlodzislaw Duch and Karol Grudziński and Grzegorz Stawski. SYMBOLIC FEATURES IN NEURAL NETWORKS. Department of Computer Methods, Nicolaus Copernicus University. [[View Context](#)].

Citation Request:

Please refer to the Machine Learning Repository's [citation policy](#)

[1] Papers were automatically harvested and associated with this data set, in collaboration with [Rexa.info](#)



In Collaboration With:



[About](#) || [Citation Policy](#) || [Donation Policy](#) || [Contact](#) || [CML](#)